

UCLA

UCLA Previously Published Works

Title

Recurrence rates provide evidence for sex-differential, familial genetic liability for autism spectrum disorders in multiplex families and twins.

Permalink

<https://escholarship.org/uc/item/6gr0b4w0>

Journal

Molecular autism, 6(1)

ISSN

2040-2392

Authors

Werling, Donna M
Geschwind, Daniel H

Publication Date

2015

DOI

10.1186/s13229-015-0004-5

Peer reviewed

RESEARCH

Open Access

Recurrence rates provide evidence for sex-differential, familial genetic liability for autism spectrum disorders in multiplex families and twins

Donna M Werling¹ and Daniel H Geschwind^{1,2,3,4,5*}

Abstract

Background: Autism spectrum disorders (ASDs) are more prevalent in males, suggesting a multiple threshold liability model in which females are, on average, protected by sex-differential mechanisms. Under this model, autistic females are predicted to carry a more penetrant risk variant load than males and to share this greater genetic liability with their siblings. However, reported ASD recurrence rates have not demonstrated significantly increased risk to siblings of affected girls. Here, we characterize recurrence patterns in multiplex families from the Autism Genetics Resource Exchange (AGRE) to determine if risk in these families follows a female protective model.

Methods: We assess recurrence rates and quantitative traits in full siblings from 1,120 multiplex nuclear families and concordance rates in 305 twin pairs from AGRE. We consider the first two affected children per family, and one randomly selected autistic twin per pair, as probands. We then compare recurrence rates and phenotypes between males and females and between twin pairs or families with at least one female proband (female-containing (FC)) versus those with only male probands (male-only (MO)).

Results: Among children born after two probands, we observe significantly higher recurrence in males (47.5%) than in females (21.1%; relative risk, $RR = 2.25$; adjusted $P = 6.22 \times 10^{-8}$) and in siblings of female (44.3%) versus siblings of male probands (30.4%; $RR = 1.46$; adj. $P = 0.036$). This sex-differential recurrence is also robust in dizygotic twin pairs (males = 61.5%, females = 19.1%; $RR = 3.23$; adj. $P = 7.66 \times 10^{-9}$). Additionally, we find a significant negative relationship between interbirth interval and ASD recurrence that is driven by children in MO families.

Conclusions: By classifying families as MO or FC using two probands instead of one, we observe significant recurrence rate differences between families harboring sex-differential familial liability. However, a significant sex difference in risk to children within FC families suggests that female protective mechanisms are still operative in families carrying high genetic risk loads. Furthermore, the male-specific relationship between shorter interbirth intervals and increased ASD risk is consistent with a potentially greater contribution from environmental factors in males versus higher genetic risk in affected females and their families. Understanding the mechanisms driving these sex-differential risk profiles will be useful for treatment development and prevention.

Keywords: Female protective model, Sex differences, Multiplex families, AGRE, Recurrence risk, Interbirth interval

* Correspondence: dhg@mednet.ucla.edu

¹Semel Institute for Neuroscience and Human Behavior, David Geffen School of Medicine, University of California, Los Angeles, 760 Westwood Plaza, Los Angeles, CA 90095, USA

²Neurogenetics Program and Department of Neurology, David Geffen School of Medicine, University of California, Los Angeles, 695 Charles E. Young Dr. South, Los Angeles, CA 90095, USA

Full list of author information is available at the end of the article

Background

Autism spectrum disorders are developmental disorders that appear early in life and are defined by impairments in social skills and language abilities, as well as restricted interests and repetitive behaviors [1]. These symptoms present heterogeneously, with some autistic children showing severe intellectual disability and poor basic daily living skills, and others with high intelligence and capacity for independence. Overall, current prevalence estimates for autism spectrum disorder (ASD) are approaching 1.5%, at 1 in 68 children, in the United States [2], an increase that is thought to be attributed to increased awareness among parents, physicians, and teachers that leads to more frequent diagnoses. ASD diagnoses are approximately four times more frequent in males than in females, and the mechanisms responsible for this sex difference are not well understood.

Genetic variation contributes strongly to ASD etiology, as evidenced by high concordance rates between twins [3,4] and high recurrence in siblings [5,6], as compared with risk in the general population. ASDs also often appear sporadically as a result of deleterious *de novo* variants that arise in a parent's germ line. The identification of these rare, *de novo* copy number and single nucleotide variants (CNV, SNV) in ASD cases from simplex families has proven especially fruitful for risk gene discovery in recent years [7–14]. However, the heritable or familial component of ASD's genetic risk architecture, although likely to account for more than 50% of genetic risk [15,16], is still poorly understood; family-based genetic linkage and association studies have identified very few replicable risk loci [17–26].

Modeling patterns of inheritance have led to the proposal that families with autistic children fall into two risk classes for ASD: a majority of low-risk families in which *de novo* variants are the primary genetic cause of ASD and a minority of high-risk families in which inherited variants follow a dominant transmission pattern for males, with reduced penetrance in females [27,28]. Sibling recurrence risk estimates from multiplex families and from an infant sibling study fit this model, finding ASD recurrence in close to 50% of later-born male children in these families [6,28]. These studies report far lower risk to later-born daughters from multiplex families (approximately 20%), consistent with the effects of a biological mechanism that protects females from manifesting an ASD phenotype.

This model of relative resilience has been termed as the female protective model, a variation on the multiple threshold liability model for ASD risk [29]. According to this model, genetic liability is distributed in the population, and males and females have different thresholds, or minimum variant loads, at which they present an ASD phenotype [30,31]. Following from the model, one would

expect affected females to carry a greater risk variant load than affected males on average, and that this variant load, if inherited as opposed to *de novo*, should be shared among siblings. This phenomenon of higher recurrence risk to family members of the less frequently affected sex is referred to as the 'Carter effect' [32]. Recent evidence of higher scores on a quantitative measure of autistic traits in siblings of female probands as compared with siblings of male probands in two large, population-based samples supports this hypothesis [33].

At the genetic level, earlier work also observed trends toward higher rates of deleterious CNVs and SNVs among autistic females [7–9,12–14]. A more recent study found a significantly higher rate of CNV and SNV risk variants in females with ASD and other neurodevelopmental disorders, as well as preferential transmission of these CNVs from mothers [34]. However, most genetic studies of rare variants have focused on collections of sporadic ASD cases, which are assumed to show enrichment for *de novo* mutation events (though it is likely that inherited genetic variants contribute to risk among these families as well). It is not known if the effects of increased variant loads in females and their siblings are also evident among multiplex families, which are assumed to be enriched for inherited risk variants, in the population or in research collections such as the Autism Genetics Resource Exchange (AGRE) cohort.

AGRE is a collection of pedigree and phenotypic data and genetic material from families with autistic children [35]. Due to an early focus on recruitment of families with multiple affected children, AGRE families have been widely utilized in genetic studies aiming to identify shared, familial risk variants, including linkage analyses [18,20–22,24–26] and family-based association testing [17,23]. The female protective model predicts that prioritizing families with affected females for variant discovery work may enrich study samples for more deleterious and detectable variants with larger effect sizes than the average familial risk variant load. Additionally, understanding the actions of female protective factors could serve to identify new therapeutic avenues.

We analyzed recurrence and concordance patterns in >1,000 multiplex families and >300 twin pairs from the AGRE cohort to test two primary hypotheses based on the female protective model: first, that males will show higher rates of ASD than females, and second, that risk will be greater for the siblings and co-twins of autistic females than siblings and co-twins of autistic males. We additionally posit several related, secondary hypotheses. One is that sex-specific risk and/or protective factors and familial genetic risk variant loads contribute simultaneously to individuals' total liability for ASD [31], such that: 1) even within families carrying high, female-penetrant genetic liability, females will show lesser risk

than males, and 2) male siblings of female probands will show higher risk than male siblings of exclusively male probands. Finally, if we extend the concept of ASD liability to include environmental risk factors, we hypothesize that siblings of autistic males, who are predicted to carry relatively lesser familial genetic liability than siblings of autistic females, will show a stronger relationship between risk-associated environmental factors and the likelihood of ASD diagnosis. If female-specific factors protect against these environmental exposures (as well as against genetic insults), then one can predict that this association between environment and ASD risk will be especially strong for the male siblings of autistic females.

Methods

Subjects

AGRE is a collection of phenotypic and genetic data from families with autistic children that was established in 1997 [35,36]. Initially founded as a multiplex cohort, AGRE currently also includes simplex families, though it remains a valuable source of multiplex ASD families for study. All subjects in AGRE provided written informed consent or assent with parental agreement for behavioral evaluation, blood sample collection, and the transfer of collected data to the AGRE program. This study was approved by the Western Institutional Review Board (AGRE), the Institutional Review Board at Washington University (subject recruitment, principal investigator: John Constantino), and by the Medical Institutional Review Board 3 at the University of California, Los Angeles.

Starting from the catalog of all AGRE subjects (database queried on 14 April 2014), which included 12,260 individuals from 2,278 families, we filtered families for inclusion in these analyses (Additional file 1: Figure S1). Extended families were first parsed to nuclear families, and in order to enrich this sample for cases with genetic risk factors as opposed to environmental complications, nuclear families that included a child with pre- or perinatal insults, or premature birth before 35 weeks, were removed. Families with twin pairs or multiples of unknown zygosity were also excluded.

For this study, we classified as affected all subjects with study diagnoses of autism, 'broad-spectrum,' or 'not quite autism' based on a clinician's evaluation of Autism Diagnostic Interview-Revised and Autism Diagnosis Observation Schedule scores. A 'broad-spectrum' diagnosis is given to individuals with pervasive developmental disorders of varying severity and includes subjects with conditions formerly termed as pervasive developmental disorder-not otherwise specified (PDD-NOS) and Asperger's syndrome. A diagnosis of 'not quite autism' is given to subjects who meet the autism cutoffs in all symptom

domains but who do not meet the age of onset criterion or conversely who meet the age of onset criterion but fall only one point short of autism cutoffs in one or more symptom domains. Families with one or more children with ambiguous diagnoses, in which AGRE clinicians did not evaluate a child but their parent reported a diagnosis from a community professional, were removed from the analysis. All monozygotic (MZ) multiples and dizygotic (DZ) twin pairs from families meeting the above criteria that included at least one affected child were included in concordance analyses.

For analyses of siblings from multiplex families, we applied additional filters. First, since genetic risk variants carried by MZ multiples are non-independent, we selected one individual from MZ sets at random for inclusion. Families with only one remaining affected child were then excluded, as were families in which affected children were half siblings of one another. The birth order of all full sibling children was then assigned by sorting the mother's or father's age at time of birth, if known. For the 12 families who lacked parental age information, birth order was assigned by sorting the individual subject identification numbers, which are assigned according to the birth order. The final multiplex sample consisted of 5,328 individuals from 1,120 nuclear families, including 2,404 affected children, 684 unaffected full siblings, and 2,240 parents.

Each multiplex nuclear family was then classified by the sex of the first two affected children born in the family (probands) either as FC with at least one affected female proband or as MO with only affected male probands. While this approach will misclassify some families with later-born autistic daughters as 'male-only,' it prevents the artifactual inflation of recurrence rates in FC families (and deflation in MO families) that results from calculating recurrence rates in the same, later-born children that are also considered during family classification. In other words, if the sex of all affected children in a family is considered during family classification, then later-born affected girls will always contribute positively to the FC recurrence rate, while all female children in MO families will be unaffected by definition, thus reducing the apparent recurrence in MO families. Using the sex of the first two affected children born to classify each family therefore allows us to more definitively separate MO from FC families than is possible from a single proband but does not systematically bias the recurrence rates that we observe in children born after the probands.

Sex ratios

We calculated the ratio of males to females from all affected children in the multiplex family set. Then, since previous studies have shown differences in the relative numbers of affected males and females among high- and

low-functioning cases [37,38], we also calculated sex ratios within the subsets of affected children who met the criteria for the strict autism diagnosis, children with lesser diagnoses of broad-spectrum or not quite autism, children with a Vineland Adaptive Behavior Scales (VABS) composite standard score within the top quartile in the sample (score ≥ 75), and children with a VABS score within the bottom quartile (score ≤ 50) [39]. We applied the VABS as the main measure of interest here as it is the most completely ascertained phenotypic measure of intellectual ability or general functioning in AGRE, with 1,656 of 2,404 (69%) cases with recorded scores. Though we do utilize the full range of recorded VABS scores for an additional assessment of quantitative phenotypes, here, we simply use the top and bottom quartiles of the VABS scores as a proxy for the most high- and most low-functioning cases within this data set.

Recurrence risk

To determine if multiplex families from AGRE show evidence of a female protective effect for ASD penetrance, we assessed ASD prevalence in siblings beyond the two affected children required per family to meet criteria for multiplex status. We then tested whether this risk differs by the sex of the evaluated children or by the families' classification as MO or FC.

In all families with additional children born after their second affected child, we assessed recurrence risk. First, we recorded the affection status of all children born after the second affected child ($N = 456$ children from 341 families); this provided the largest available sample of latter-born children for estimating recurrence rates. Next, we estimated recurrence in three variations of this sample in order to give equal weight to each nuclear family regardless of size and to directly replicate and extend previous analyses of recurrence in the AGRE sample.

In the first variation, we recorded the affection status of only the first child born after the second affected ($N = 341$ families); this method ensures that all families contribute independently to the risk calculation but limits sample size. In the second variation, we replicated the method applied in a 2007 study of recurrence risk in AGRE families [28] by recording the affection status of the third child from families with exactly three children in which the first two children are affected with ASD ($N = 198$ families). This strict approach was applied to control for effects of 'stoppage,' or parents' decisions to curtail their intended family size after having children with ASDs, on recurrence risk estimates. Here, for the third variation, we also extended this strategy to test the last-born child in all families who had only one additional child after their second affected, regardless of the total family size or birth order of the first two affected children. This extension allowed us to include a greater

number of families ($N = 258$ families) than that used in the method from Zhao *et al.* [28] while still controlling for potential stoppage effects.

To evaluate the risk across multiplex families without limiting this analysis to later births in families who continued having children and without weighting these estimates by including multiple children from large families, we next calculated what we refer to as 'familial risk.' The main purpose of this analysis was to determine whether the patterns we observe in sample variations of latter-born children are also reflected in the larger sample of multiplex families from AGRE. In all families with at least three children, we calculated 'familial risk' by running 1,000 trials in which two affected children are masked at random and affection status is evaluated in a third child. Per trial, the sexes of the masked affected are also used to classify the family as FC or MO. Familial risk is taken as the mean risk from these 1,000 randomizations. This method allowed for the inclusion of those families who stopped having children after their second affected, as well as all children in each family regardless of birth order. Since only one child per family was considered in each randomized trial, this approach also ensured that large families would not contribute disproportionately to the risk estimate.

For each of these five estimates of recurrence or familial risk - (A) all subsequent children from all families, (B) the single next-born child from each family, (C) the third-born child from three-child families, (D) the last-born child from families where the second affected child is born second to last, and (E) familial risk from 1,000 random selections of one child per family - we performed one-sided Fisher's exact tests in JMP (SAS Institute, Inc., Cary, NC, USA) to compare the risk in males to females and in FC to MO families. We also compared the males' and females' risk within FC families and males' risk in FC to males' risk in MO families. *P* values were adjusted for these 20 tests by Bonferroni correction. Additionally, we applied logistic regression models to test for interaction effects of sex by family type.

Birth order and interbirth interval

Earlier work has reported increased ASD risk for children born shortly after elder siblings in population samples [40-42]. It is not known how a risk factor such as short interbirth interval (IBI) interacts with genetic risk profile. Additionally, though one might assume that risk for ASD is constant across the births in a family who share a common source of heritable genetic risk variation, it may be that risk for ASD is increased in later-born children due to the additional accumulation of deleterious variants in the germ line with increasing parental age [43], for example. So, we investigated the relationships between birth order and IBI on recurrence risk

in our family sample. For birth order, we assessed differences in ASD recurrence rates for children born first versus second after a family's second affected child. Risk was evaluated separately by subjects' sex and family type, and two-sided Fisher's exact tests were applied to identify significant differences in risk between birth order positions. For IBI, we used parents' age at the time of each child's birth to calculate the number of months in between siblings' births. Using logistic regression for affection status by the natural log of IBI in months, we then tested the relationship between IBI and ASD risk separately by sex and family type (MO or FC) in the child born first after the second affected child from the 332 families with complete parental age information. Maternal and paternal ages were considered as covariates. Pregnancy and birth complications were not explicitly considered in this model, since all families with any records of pre- or peri-natal insults or premature birth before 35 weeks were excluded from this and all analyses in this report.

Quantitative phenotypes

Previous studies have reported an exacerbation of the male bias for ASD among high-functioning individuals and a greater representation of females among cases with intellectual disability [37,38,44,45]. In contrast, the female protective model predicts that males, who lack female protective factors, should be more severely impacted than females by genetic risk loads of comparable magnitude; this differential impact may be detectable as shifts in phenotype severity, measured quantitatively. We tested several quantitative phenotypes related to ASD severity and intellectual ability, including the VABS composite standard score [39] (1,656 recorded scores - 69% of cases), the Peabody Picture Vocabulary Test (PPVT) standard score [46] (1,386 recorded scores = 58% of cases), the Raven's Progressive Matrices estimated non-verbal intelligence quotient (Raven's NVIQ) [47] (1,316 recorded scores = 55% of cases), and the Social Responsiveness Scale (SRS) raw total score [48] (1,042 recorded scores = 43% of cases), for sex differences overall and within FC families, and for differences between MO and FC families. Affected subjects missing scores on these measures were more likely to have been ascertained early on in AGRE's collection than more recently (as approximated by the sequentially assigned ID number for each family; difference in ID index between subjects with missing versus recorded scores for VABS = -487.93, standard error (std. err.) = 46.11, $P < 1e-04$; PPVT diff. = -385.96, std. err. = 44.47, $P < 1e-04$; Raven's NVIQ diff. = -282.98, std. err. = 44.37, $P < 1e-04$; SRS diff. = -0.563, std. err. = 43.36, $P = 0.99$) and were also more likely to have been born later in their family (difference in birth order between subjects with missing versus recorded scores for VABS =

0.047, std. err. = 0.043, $P = 0.27$; PPVT diff. = 0.24, std. err. = 0.040, $P < 1e-04$; Raven's NVIQ diff. = 0.24, std. err. = 0.039, $P < 1e-04$; SRS diff. = 0.14, std. err. = 0.038, $P = 2e-04$), than affected subjects with recorded scores.

We used scores as recorded by AGRE, and in cases where a child was evaluated more than once, we used the most recent score for analysis. All scores recorded as 'untestable' were set to missing. For the Raven's NVIQ, some children received scores of 'ATN' (above the highest possible NVIQ score normalized by age; $N = 93$) or 'BTN' (below the lowest possible NVIQ score normalized by age; $N = 19$). These scores were recoded as 160 and 20, which are above the observed maximum and below the observed minimum NVIQ scores in the remaining subjects. These high and low values match the maximum and minimum scores for the PPVT and VABS standard scores, two metrics that are scaled analogously to standard IQ.

Scores for the VABS, PPVT, and Raven's NVIQ are positively correlated with one another, with correlation coefficients of 0.447 (VABS with Raven's NVIQ), 0.545 (PPVT with Raven's NVIQ), and 0.595 (VABS with PPVT). As higher SRS scores indicate more severe ASD traits, SRS is negatively correlated with the above measures ($r = -0.291$ with Raven's NVIQ; $r = -0.395$ with PPVT; $r = -0.573$ with VABS). However, each of these instruments measures a different aspect of abilities (adaptive behavior, vocabulary ability, non-verbal intelligence) or symptoms (traits specifically associated with the ASD phenotype), and so, we opted to include all of these tests in our analyses.

Sex and family classification comparisons were assessed by *t*-tests allowing for unequal variances in JMP using the scores from one proband selected at random from each nuclear family. To test for phenotypic differences between FC and MO families that are not potentially confounded by potential sex differences in phenotypic measures, randomly selected male probands were also compared. To test for sex-differential phenotypes within FC families, a paired *t*-test was used to compare scores from one randomly selected affected female and one affected male within each family. *P* values were adjusted for 16 tests by Bonferroni correction.

Concordance in twin pairs

MZ (111 twin pairs and 1 set of quadruplets) and DZ twins (193 pairs) with at least one affected member from families without perinatal complications or ambiguous diagnoses were evaluated for ASD concordance. MZ multiples were stratified by their sex (female-female (F-F) and male-male (M-M)) and tested for concordance rate differences using the one-sided Fisher's exact test ($F-F > M-M$). For DZ twin pairs, we selected one affected twin from each pair as the proband twin, and we

compared ASD recurrence rates in the co-twin by the sex of the co-twin and the sex of the proband (analogous to FC versus MO comparison for siblings) using a one-sided Fisher's exact test (male co-twin > female co-twin and female proband > male proband). We also used a logistic regression model to test for an interaction effect of proband sex by co-twin sex on ASD recurrence rates in this sample of DZ twins.

Results

Sex ratios

Within 1,120 nuclear families with two or more full sibling children with diagnoses of ASD, there are 2,404 affected children, including 1,867 affected males and 537 affected females for an overall male-to-female ratio in these families of 3.48 (Table 1). In contrast with reports of even greater male skew among less severely affected cases [37], the sex ratio for children with AGRE diagnoses of broad-spectrum and NQA ('not quite autism') was 2.11, compared with 3.71 for children diagnosed with autism. A comparison of sex bias within cases scoring in the top and bottom quartiles from this sample on the VABS, the most completely ascertained measure of general functioning in the sample, showed a similar pattern, with a greater proportion of affected females falling in the high functioning quartile of the scale (M:F = 2.35) than the lower functioning quartile (M:F = 4.05). We note that these unexpected patterns may not accurately reflect trends at the general population level and may instead be a consequence of the multiplex ascertainment scheme for AGRE.

Recurrence risk

The recurrence rate for ASD in the multiplex set of families with at least one child born after the second affected ($N=456$ children from 341 families) was 36.0% (Additional file 2: Table S1A). The recurrence rate in male children was 47.5% and 21.1% in female children, a significant difference ($P=3.11\text{e-}09$, adj. $P=6.22\text{e-}08$), representing a male to female (M:F) relative risk (RR) of 2.25 (Figure 1A); these sex-differential rates closely match those observed by Zhao and colleagues analyzing

a sample of 165 AGRE families [28]. We also observed a difference in risk between FC and MO families, with 44.3% recurrence in FC and 30.4% recurrence in MO families ($P=1.78\text{e-}03$, adj. $P=0.036$) for a RR of 1.46 for FC families compared with MO (Figure 1B). The sex difference in recurrence risk between males and females within FC families was robust, with RR of 1.85 ($P=7.18\text{e-}04$, adj. $P=0.014$), and the difference between males from FC and MO families was nominally significant (RR = 1.27, $P=0.043$, adj. $P=0.86$; Figure 1C). When only the first child born after the second affected child was included ($N=341$ children), we find a RR of 2.36 in males compared with females ($P=1.21\text{e-}08$, adj. $P=2.42\text{e-}07$) and 1.27 in FC compared with MO families ($P=0.051$, adj. $P=1$; Additional file 2: Table S1B). Sex-differential risk was again apparent within FC families, with RR of 2.07 ($P=7.19\text{e-}04$, adj. $P=0.014$).

Since parents' decision to curtail their intended family size after having an autistic child, or stoppage, impacts analyses of family structure, we calculated recurrence risk from 198 families with a specific structure: exactly three full sibling children, with affected first and second children, as was used to estimate recurrence rates in an earlier study of AGRE families [28]. Recurrence risk patterns in this specific set were comparable to those from all families with additional births, with 38.9% risk overall, a RR of 2.39 in males versus females ($P=9.57\text{e-}06$, adj. $P=1.91\text{e-}04$), RR of 2.31 in males versus females from FC families ($P=1.86\text{e-}03$, adj. $P=0.037$), and a RR of 1.46 in FC versus MO ($P=0.025$, adj. $P=0.49$; Additional file 2: Table S1C). When this test is expanded to consider all families who had only one more child after their second affected ($N=258$ families), recurrence risk again follows the same pattern (M:F RR = 2.41, $P=6.87\text{e-}07$, adj. $P=1.37\text{e-}05$; M:F in FC families RR = 2.06, $P=3.59\text{e-}03$, adj. $P=0.072$; FC:MO RR = 1.43, $P=0.02$, adj. $P=0.39$; Additional file 2: Table S1D). The difference between males from FC versus MO families only reached nominal significance in the set of 198 families with three children (RR = 1.44, $P=0.034$, adj. $P=0.68$).

We also calculated familial risk by applying a randomization procedure that permitted inclusion of all

Table 1 Ratio of affected males to females in multiplex families from AGRE

	Number of families	Number of affected children	Number of affected males	Number of affected females	Male: Female
All diagnoses ^a	1,120	2,404	1,867	537	3.48
Autism	1,106	2,158	1,700	458	3.71
Spectrum ^b	220	246	167	79	2.11
Top quartile (≥ 75) VABS ^c	319	445	312	133	2.35
Bottom quartile (≤ 50) VABS ^c	306	424	340	84	4.05

^aAll diagnoses = autism, broad-spectrum, or not quite autism (NQA) study diagnoses; ^bSpectrum = broad-spectrum or NQA; ^cVABS quartiles calculated from composite standard scores for affected children from AGRE multiplex families. VABS, Vineland Adaptive Behavior Scales.

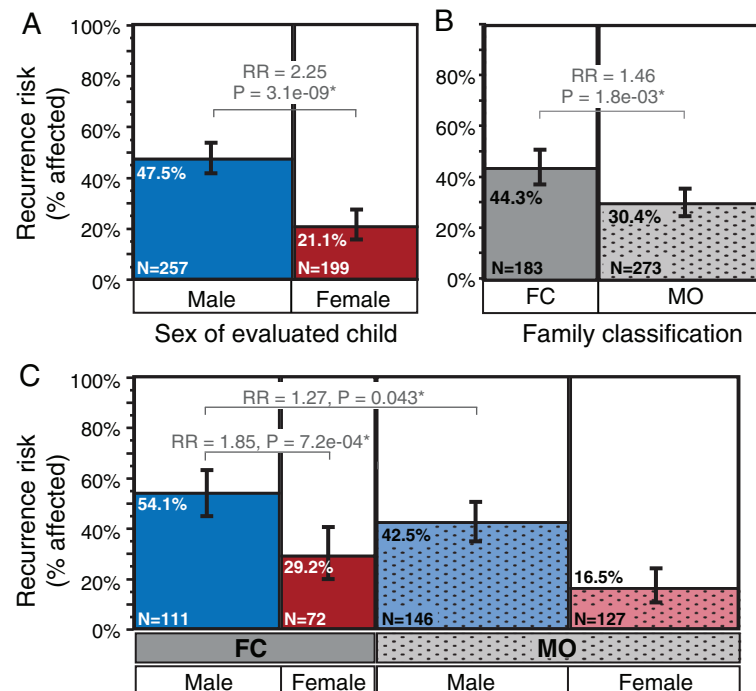


Figure 1 Recurrence risk for ASD in multiplex families from AGRE by sex and family type. ASD risk to all children born first after the second affected child in each family ($N = 456$ children from 341 families) is highest for males and in female-containing families. Mosaic plots show the proportion of affected children by (A) sex, (B) family type as FC (female-containing) or MO (male-only), and by (C) sex within each family type. Bar widths are proportional to the number of children or families in each group, which is also noted in the figure. Whiskers note the 95% confidence intervals around each recurrence rate estimate.

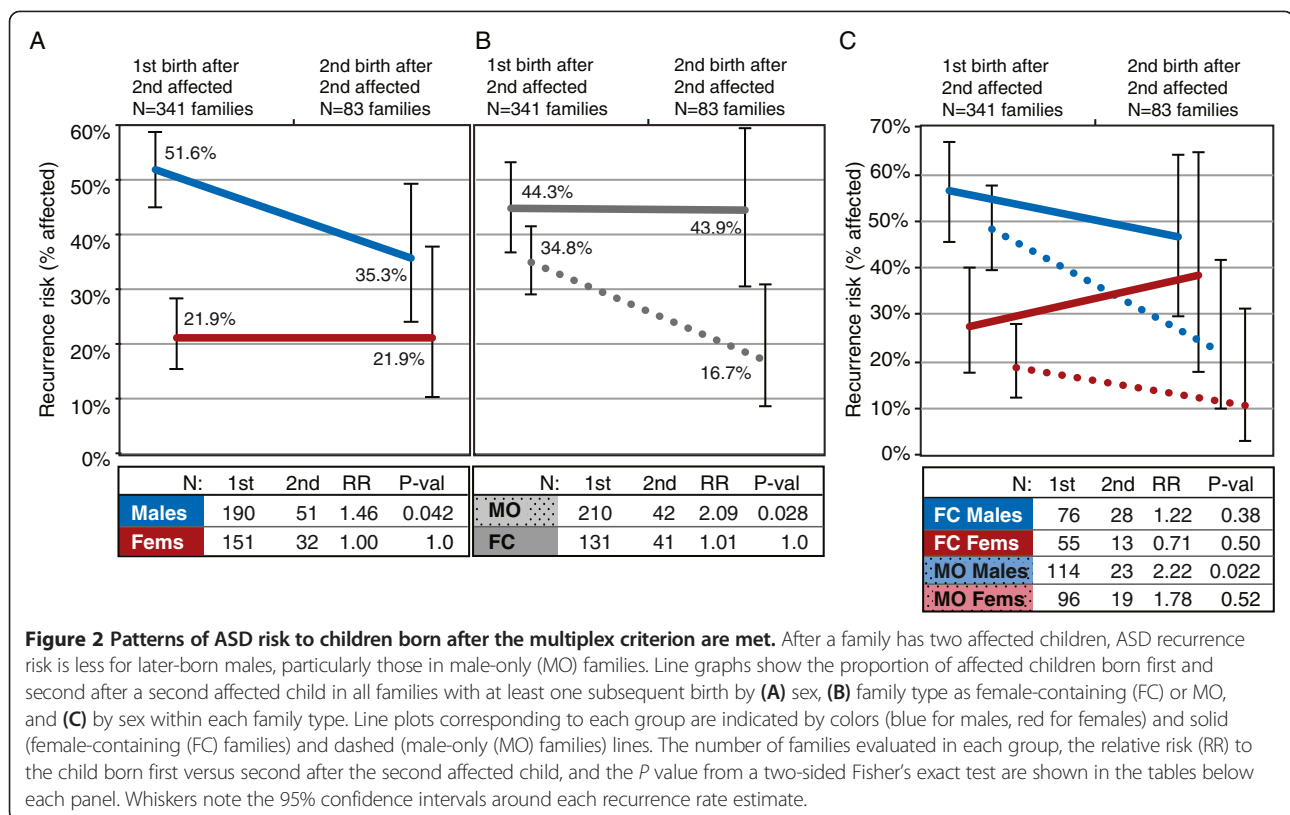
556 families with at least three full sibling children in the analysis. Familial risk in these families was found to be 17.5% (Additional file 2: Table S1E), less than the overall recurrence risk. This is likely a simple consequence of including 221 families with unaffected, earlier-born children; these families do not contribute to any recurrence rate calculations. The absolute familial risk estimates within each sex and family classification were similarly low, at 25.56% for males and 9.79% for females and 22.60% for FC and 14.37% for MO families. Relative risks showed slightly more pronounced differences than for recurrence risk (M:F $RR = 2.61$, $P = 6.70e-07$, adj. $P = 1.34e-05$; FC:MO $RR = 1.57$, $P = 9.72e-03$, adj. $P = 0.19$). Logistic regression for affection status in the later-born children from each of the family sets described above, or the non-masked, randomly selected children from familial risk calculations, additionally demonstrated significant main effects of sex and family type on risk for ASD. However, in all family sets tested, the interaction between sex and family type did not reach significance (Additional file 2: Table S2).

Birth order and interbirth interval

When comparing recurrence risk by birth order, between children born two versus one births after the second affected child, we find no significant differences in

risk to females overall and to children in FC families. However, we observe that risk to males overall and risk to children in MO families shows a trend toward a decrease from the first to the second post-affected child (males: $P = 0.042$, adj. $P = 0.33$; MO: $P = 0.028$, adj. $P = 0.23$; see Figure 2A,B). When children are stratified by both sex and family type, we find that the risk to either males or females from FC families does not differ significantly between the first and second post-affected children. In contrast, risk to males in MO families drops for the second post-affected child, from 48.25% to 21.7% ($P = 0.022$, adj. $P = 0.18$; Figure 2C). These patterns are comparable when considering only the 83 families with at least two children born after their second affected child (Additional file 1: Figure S2A-C).

We also find a significant negative association between the number of months since the birth of the second affected child (IBI) and ASD risk to the next-born child ($\chi^2 = 10.41$, $P = 1.25e-03$; Table 2). Within subgroups of children, this effect is significant for males overall ($P = 2.31e-04$; Figure 3A), for children in MO families ($P = 6.63e-03$; Figure 3B), and for male children in MO families ($P = 5.1e-04$; Figure 3C). The relationship between IBI and ASD status does not reach significance for any subgroup of females or children from FC families, consistent with the hypothesized existence of a



maternal, uterine risk mechanism that predominantly affects MO families. However, when sex and family type (FC or MO) are both included as factors in the regression model for recurrence, only the main effects sex and IBI, not family type, are significant (Additional file 2: Table S3). We also note that the significance of these effects appears to be driven by unaffected children with

long IBIs (Figure 3D,E); when we evaluate only those children born within 60 months of the second affected in their family, the associations between IBI and ASD risk diminish (males $P = 0.05$, females $P = 0.600$, FC $P = 0.6$, MO $P = 0.4$).

Quantitative phenotypes

To determine if males and females or the different family types differed from one another in the presentation of ASD and its impact on functioning, we assessed quantitative measures of ASD severity, general functioning, and intellectual ability. To compensate for the non-independence of multiple children from each family, we compared VABS, PPVT, Raven NVIQ, and SRS scores from one randomly selected proband per family. We observed significantly higher VABS (better adaptive functioning) scores in probands from FC compared with MO families (average FC-MO difference = 4.42, $P = 2.60 \times 10^{-3}$, adj. $P = 0.042$). This difference was also nominally significant when comparing only male probands (FC-MO difference = 2.85, $P = 0.049$). We also observed significantly lower VABS scores in males from FC families as compared with their sisters by a paired test (average difference = -2.85, $P = 0.035$). No comparison of scores from the PPVT, Raven's NVIQ, and SRS showed any significant group differences on these measures.

Table 2 Autism recurrence risk by interbirth interval

Group	N	Chi-squared	P value	Adjusted P value
All	332	10.41	1.25e-03	0.011*
Males	188	13.56	2.31e-04	2.08e-03*
Females	144	0.52	0.471	1
FC	129	2.7	0.100	0.903
MO	203	7.37	6.63e-03	0.060
FC, Males	75	1.69	0.194	1
FC, Fems	54	0.67	0.413	1
MO, Males	113	12.08	5.10e-04	4.59e-03*
MO, Fems	90	0.049	0.825	1

Chi-squared statistics and *P* values are from the whole-model test for logistic regression of affection status of the first child born after the second affected child by the natural log of the interbirth interval (time since the birth of the second affected child) in months. Adjusted *P* values have been corrected for nine tests. *Adjusted *P* value ≤ 0.05 . MO, male-only families; FC, female-containing families.

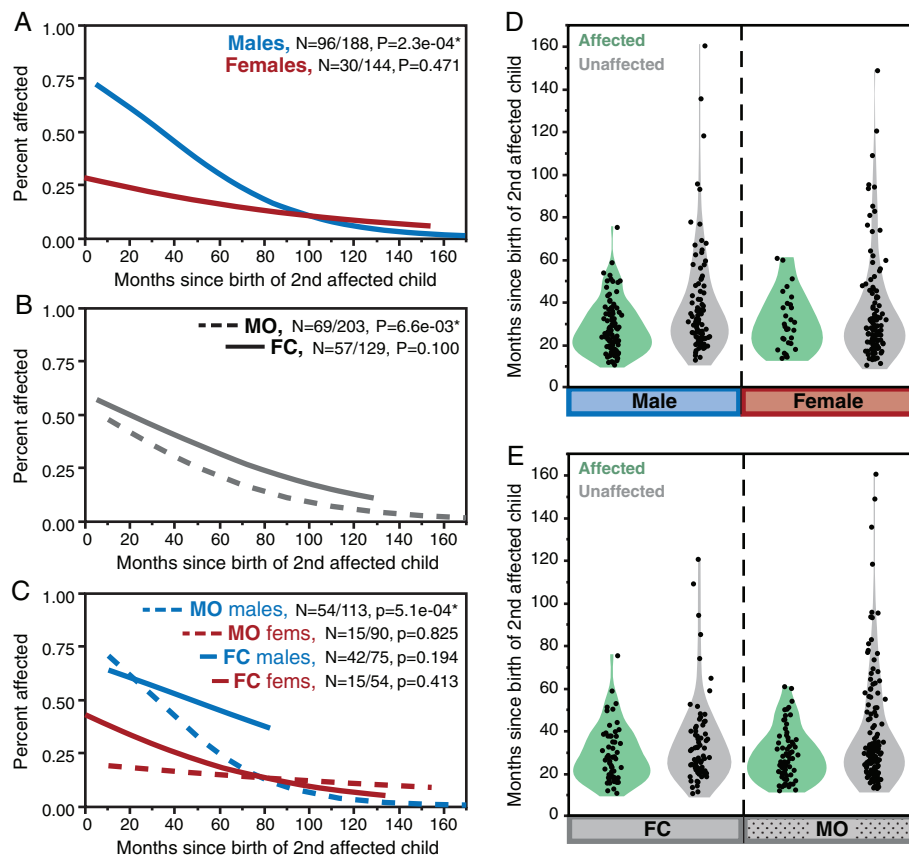


Figure 3 Relationship between interbirth interval and ASD recurrence risk by sex and family type. Recurrence risk to subsequent children decreases with increasing interval between births. (A), (B), and (C) show the logistic regression estimated recurrence risk by the number of months since the birth of the second affected child within each subgroup of individuals as indicated by colors (blue for males, red for females) and solid (female-containing (FC) families) and dashed (male-only (MO) families) lines. P values from the whole-model test of logistic regression for affection status given the natural log of interbirth interval (months) are noted; fractions show the number of affected individuals out of the total within the indicated subgroup. (D, E) show the interbirth interval for all evaluated children given their affection status and either their sex (D) or family type (E).

Concordance in twin pairs

We assessed concordance rates in 112 MZ multiples and 193 DZ twin pairs from AGRE. We identified high concordance rates in MZ multiples with male pairs (M-M) showing 95.6% concordance and female pairs (F-F) showing 85% concordance (F-F:M-M $RR = 0.89$, $P = 0.11$; Figure 4A; Additional file 2: Table S4A). For DZ twin pairs, we find a significantly higher recurrence rate among male co-twins than female co-twins (61.5% versus 19.1%; M:F $RR = 3.23$, $P = 1.92e-09$, adj. $P = 7.66e-09$; Figure 4B; Additional file 2: Table S4B); this sex difference is also apparent when co-twins of female probands are tested separately (71.4% versus 20.0%; M:F $RR = 3.57$, $P = 2.97e-03$, adj. $P = 0.012$; Figure 4D; Additional file 2: Table S4B). We also observe a trend toward higher recurrence rate in co-twins of female probands than male probands (50.0% versus 41.4%; F-pro:M-pro $RR = 1.21$, $P = 0.23$; Figure 4C; Additional file 2: Table S4B). Logistic regression for co-twin affection status corroborated these results by

demonstrating only a significant main effect of co-twin sex ($P = 3.45e-06$); neither the proband sex nor the interaction (proband sex by co-twin sex) terms were significant in this model (Additional file 2: Table S4C).

Discussion

The first main hypothesis derived from the female protective model is simply that males should demonstrate greater ASD risk than females. Our findings confirm this prediction, as we observe 2.25- to 2.6-fold increased recurrence risk to males compared with females. This increased risk for ASD in males compared with females corroborates findings from other studies of recurrence risk in infant siblings [6] as well as in a sample of families from AGRE and the Interactive Autism Network registry [5]. The observed recurrence rates of approximately 50% in males and 20% in females are also entirely consistent with those reported by Zhao *et al.* [28], whose earlier analysis also utilized a smaller sample of AGRE

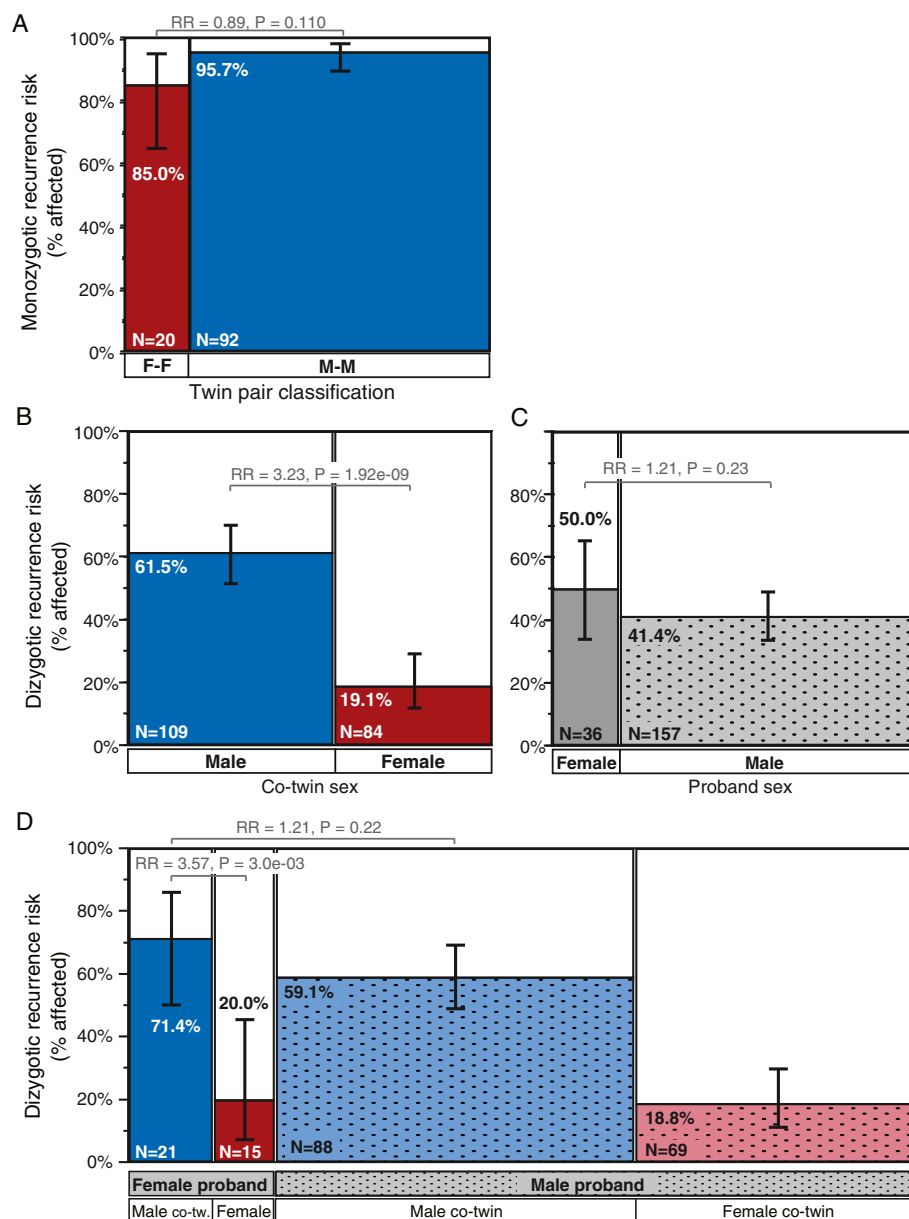


Figure 4 Concordance rates in monozygotic and dizygotic twin pairs classified by their sex. ASD concordance rates are higher for monozygotic than dizygotic twins, and concordance in dizygotic twins shows an effect of sex. **(A)** Mosaic plot shows the proportion of concordant monozygotic (MZ) twin pairs by the sex of the pair (F-F = female-female, M-M = male-male). Mosaic plots in **(B)**, **(C)**, and **(D)** show recurrence rates for ASD in co-twins from dizygotic (DZ) twin pairs by co-twin sex **(B)**, proband twin sex **(C)**, and by each combination of proband and co-twin sex **(D)**. Bar widths are proportional to the number of twin pairs in each group, which is also noted in the figure, and whiskers note the 95% confidence intervals around each recurrence risk estimate. Relative risk (RR) and *P* values for each comparison of interest are also noted in the figure.

families available at the time. Additionally, recurrence rates in DZ twins confirm this hypothesis as well, with male co-twins showing a 3.23-fold increased risk compared with female co-twins.

The second primary hypothesis from the female protective model is that the siblings of female probands should demonstrate greater ASD risk than the siblings of male probands, which we observe. We note that only

when all children born after the second affected child are included in the recurrence rate calculation does this comparison survive a conservative correction for multiple testing. In other family structures (last child from three-child families, last child from families of three or more, familial risk estimation), where only one child per family is permitted to contribute to the recurrence risk estimate, the differences between FC and MO families

no longer reach significance. These shifts in statistical significance between family sets utilizing one or more later-born affected children in the recurrence risk estimate suggest that this analysis is likely to be underpowered. The findings from the analysis of 193 DZ twin pairs suggest a similar issue, as DZ twin pairs show a non-significant trend toward higher recurrence risk in co-twins of female than of male probands.

We also acknowledge that the maximum magnitude of the Carter effect that we are able to observe in these data is limited by our stratification approach, which considers the sex of only two, early-born probands per family, or one proband twin per twin pair. We apply this approach to avoid systematic increases in risk estimates for FC families and twins that result from using the same children to both stratify families and to calculate family-stratified recurrence rates. By the proband-based approach that we use, families that are found to carry female-penetrant risk loads *post hoc*, as evidenced by affected, later-born female children or affected female co-twins, may be grouped with families and twin pairs of exclusively male affected children for analysis. If these 'misclassified' families and twin pairs carry the sorts of high genetic risk loads that are responsible for the Carter effect, this will increase recurrence rate estimates among 'MO' families and attenuate the observable difference between FC and MO families and twins. Therefore, the observation of higher recurrence in imperfectly separated FC families or twins compared with MO that overcomes this counter-hypothetical skew introduced by our method can be interpreted as robust evidence of the Carter effect.

Our secondary hypotheses address the relationship between subject sex and familial genetic risk load. We predicted that females from FC families should show reduced risk compared with males from these same families, which is observed. Recurrence rates in later-born female children from FC families are significantly lower than for male children from these families and are also reduced in female co-twins of female probands, demonstrating the impact of female-protective mechanisms even within highly risk-loaded FC families and twins. The observation of a consistent direction of recurrence rate differences between FC and MO males and the differences that we observe when both sexes are considered suggest that the lack of a significant difference among males from different family types is likely to be a consequence of low power within this subset of samples. Analysis of multiplex families from other, larger collections will be necessary to conclusively reject the null for the impact of sex-differential familial liability on ASD risk in males.

Alternatively, the lack of a significant increase in risk to FC males over MO males may suggest a potential modification to the multiple threshold liability model.

As opposed to conceptualizing genetic risk load severity as purely quantitative, there may be a locus-specific component such that females are only vulnerable to the effects of a subset of specific inherited risk variants that uniformly increase males' risk. Risk variants may be assigned to one of the two classes: 1) variants that are penetrant in both males and females and 2) variants that are predominantly penetrant in males. These variant classes may be functionally delineated by the specific loci that harbor the risk variants, in that some loci increase ASD risk in both sexes (relative to sex-differential, baseline population risk) but females are nearly fully protected from variants at other loci. There are a handful of risk loci such as SHANK1 microdeletions [49] and 16p13.11 CNVs [50] that have been reported to follow such a pattern, where male carriers manifest ASD or other neurodevelopmental conditions and female carriers do not.

Of the four quantitative phenotypic measures of intellectual ability tested and for all comparisons of interest, only the VABS showed post-correction and nominally significant score differences. In agreement with our hypotheses, paired tests comparing VABS scores in male versus female siblings from FC families show nominally lower scores in brothers compared with their sisters, consistent with the prediction that males would be more severely impacted by female-penetrant risk loads. These data suggest that genetic and sex-differential risk loads in these families impact liability for the ASD phenotype, but that they have minimal to no consistent impact on the measureable degree of symptom severity, intellectual ability, or adaptive functioning.

Lastly, we posited that siblings of autistic males should show a stronger contribution from non-genetic, environmental risk factors on their liability for ASD than siblings of autistic females, whose risk is predicted to be more completely derived from a larger genetic liability. The finding that IBI is a significant predictor of ASD recurrence risk supports this hypothesis. In agreement with previous studies of population-based cohorts from California [40,42] and Norway [41], we observe a negative relationship between IBI and ASD risk, with no autistic cases born more than 75 months after their next eldest, autistic sibling and with children born after short IBIs showing the highest recurrence rates. We find that IBI is a significant predictor of affection status only for males, specifically those born into MO families. This finding is consistent with the idea of males as more vulnerable to risk factors in general, as well as with our hypothesis that lesser variant loads in MO families may leave room for contributions from non-genetic risk factors. Such non-genetic factors may include events in uterine or early postnatal development; maternal stress, inflammation, and deficiency of micronutrients such as

folic acid have all been hypothesized as potential causes of the increased risk for ASD in children born after short IBIs [41,42,51]. Discordance patterns in DZ twins are also consistent with this concept of non-heritable risk factors from which females are protected, as female DZ twins are far less likely to be affected with ASD than their male co-twins. Currently, the roles of these proposed factors in ASD risk remain speculative. Future work is needed to definitively identify the potential maternal factors involved.

Previous studies report recurrence risks of 10% to 20% overall, far lower than observed here. These lower frequencies can most likely be attributed to the joint consideration of families with both inherited and *de novo* genetic risk architecture, as approximated by multiplex and simplex family structures, respectively. Supporting this, study of high-risk infant siblings reported greater recurrence risk in the subset of families with two or more elder affected of 32.2% overall and nearly 50% in males [6]; these rates are entirely consistent with what we observe here in AGRE.

Additionally, several studies of larger samples and population-based cohorts have tested and failed to observe a significant effect of older affected siblings' sex on risk to later-born children [5,6,52,53]. One major difference between these study designs and that applied here is that we utilize two probands per family to classify families as FC or MO. By considering a greater number of affected children from each family, we are able to achieve a cleaner delineation between families with female-penetrant and male-specific risk loads. Also, the birth of an affected female child at any time in a family's pedigree (or here within the first two affected, for methodological reasons) serves as a positive indicator of a high familial liability load that is likely to have a larger effect size as compared with the heritable variants carried by MO families on average. Therefore, genetic studies that focus on FC families may have increased power to detect and implicate heritable risk variants, which have so far remained largely elusive.

Previous work has reported closer to equal representation of autistic males and females among severely impacted cases with comorbid intellectual disability and a more pronounced male bias among high-functioning individuals [37,38]. We observe the opposite pattern, which may be characteristic of multiplex families or specific to AGRE. With regard to the reports from other samples on sex ratios in high- and low-functioning individuals, it has been suggested that current diagnostic tools are calibrated to a male-typical phenotype and that females are under-diagnosed for ASD due to their non-prototypical presentation of ASD symptoms [54,55]. For females who lack comorbid intellectual disability, diagnoses may be especially elusive. However, in this AGRE

sample, the increased number of higher-functioning females may be a consequence of the ascertainment of families with multiple diagnosed children, and/or it may be that parents with an autistic child are more perceptive of symptoms in their daughters regardless of her intellectual ability.

The increased number of diagnosed females relative to males in this sample as compared with the general population (male/female relative risk of approximately 2.25 versus 4.5 in population samples from the United States [2]) may also be a consequence of an increased sensitivity to females' symptoms. Findings from studies that are designed to evaluate all female and male subjects equivalently (as opposed to analyzing existing diagnostic records) support this possibility, including recent epidemiological population screens (South Korea, M:F = 2.5:1 [56]; Finland, M:F = 2:1 [57]) and a study of infant siblings of autistic probands (M:F = 1.65:1 [58]). It is also possible that male, later-born children in multiple incidence families have less ASD risk, or that females have greater ASD risk, than their earlier-born siblings and that this change drives the attenuated relative risk that we observe in our sample. Here, we do observe a decline in male children's recurrence rates between the first and second births after a second affected child (but no change in female children's recurrence), consistent with such a sex-differential birth order effect on risk. Though here, this birth order effect was only a trend and thus warrants further investigation.

We also comment that the burden of care required by an autistic child can be substantial, and so, in addition to potential differences at the genetic risk level, there may be key differences between those parents who continue to have children after their earlier-born child is diagnosed with ASD and parents who do not. Though they may carry highly penetrant, heritable risk variants, the latter families will appear as simplex and are therefore not characterized here. A comparison of quantitative phenotypic measures in families who stopped versus continued having children after their second affected child does show trends of lower VABS scores in families who stopped, though these differences do not reach significance after adjustment for multiple testing (Additional file 2: Table S6).

Conclusions

We characterize recurrence risk in a large cohort of uniformly ascertained multiplex families and observe significant sex differences in recurrence rates, with females showing a greater than twofold reduction in risk compared with males' risk. We also observe higher recurrence rates in families with at least one affected female proband as compared with families whose probands are exclusively male; this difference is expected under the

sex-differential threshold liability model for ASD. We further report a significant relationship between ASD and interbirth interval that is driven by male siblings of male probands. Taken together, these observations demonstrate that sex, genetic risk load, and putative environmental exposures all contribute to liability for ASD and that families with autistic females comprise a set of multiplex families enriched for larger genetic risk variant loads. Identification of female protective mechanisms in these high-risk families would open up new therapeutic windows, and the work to discover sex-differential heritable risk variant loads will advance our understanding of the shared familial component of ASD risk.

Additional files

Additional file 1: Figures S1 and S2. **Figure S1.** Filtering applied to select 1,120 multiplex families for analysis. **Figure S2.** Patterns of ASD risk to children born after the multiplex criterion are met in families with at least two subsequent births.

Additional file 2: Tables S1 to S6. **Table S1.** Recurrence risk in multiplex AGRE families by sex and family type. **Table S2.** Logistic regression for affection status by sex and family type. **Table S3.** Logistic regression for affection status by interbirth interval. **Table S4.** Concordance rates in monozygotic and dizygotic twin pairs. **Table S5.** Adaptive behavior, intellectual ability, and ASD symptoms by sex and family type. **Table S6.** Adaptive behavior, intellectual ability, and ASD symptoms by family stoppage status.

Abbreviations

AGRE: Autism Genetics Resource Exchange; ASD: autism spectrum disorder; ATN: above the highest possible score for Raven's Progressive Matrices normalized by age; BTN: below the lowest possible score for Raven's Progressive Matrices normalized by age; CNV: copy number variant; DZ: dizygotic; FC: female-containing; IBI: interbirth interval; IPI: interpregnancy interval; IQ: intelligence quotient; M:F: male-to-female ratio; MO: male-only; MZ: monozygotic; NVIQ: non-verbal IQ; PDD-NOS: pervasive developmental disorder-not otherwise specified; PPVT: Peabody Picture Vocabulary Test; RR: relative risk; SNV: single nucleotide variant; SRS: Social Responsiveness Scale; VABS: Vineland Adaptive Behavior Scales.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

DMW conceived of the study, performed all the data processing and analyses, and drafted the manuscript. DHG assisted with the study design, interpretation, and manuscript preparation. Both authors read and approved the final manuscript.

Acknowledgements

We gratefully acknowledge resources provided by the Autism Genetics Resource Exchange (AGRE) collection and the participating families, as well as Clara M. Lajonchere and Eve Landa. Thank you also to Jennifer Lowe for the assistance with data acquisition and management. AGRE is a program of Autism Speaks and is supported, in part, by grant 1U24MH081810 from the National Institute of Mental Health (NIMH) to Clara M. Lajonchere. DMW was supported by the NIMH F31 MH093086. This work was additionally supported by ACE Network grant R01 MH100027 and ACE Center grant R01 MH071425 to DHG.

Author details

¹Semel Institute for Neuroscience and Human Behavior, David Geffen School of Medicine, University of California, Los Angeles, 760 Westwood Plaza, Los Angeles, CA 90095, USA. ²Neurogenetics Program and Department of

Neurology, David Geffen School of Medicine, University of California, Los Angeles, 695 Charles E. Young Dr. South, Los Angeles, CA 90095, USA.

³Center for Autism Research and Treatment, Semel Institute, David Geffen School of Medicine, University of California, Los Angeles, 760 Westwood Plaza, Los Angeles, CA 90095, USA. ⁴Center for Neurobehavioral Genetics, Semel Institute, David Geffen School of Medicine, University of California, Los Angeles, 695 Charles E. Young Dr. South, Los Angeles, CA 90095, USA.

⁵Department of Human Genetics, David Geffen School of Medicine, University of California, Los Angeles, 695 Charles E. Young Dr. South, Los Angeles, CA 90095, USA.

Received: 19 September 2014 Accepted: 29 January 2015

Published online: 13 May 2015

References

1. American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 5th ed. Arlington, VA: American Psychiatric Association; 2013.
2. Developmental Disabilities Monitoring Network Surveillance Year 2010 Principal Investigators CfDCaP. Prevalence of autism spectrum disorder among children aged 8 years - autism and developmental disabilities monitoring network, 11 sites, United States, 2010. *MMWR Surveill Summ*. 2014;63(2):1–21.
3. Bailey A, Le Couteur A, Gottesman I, Bolton P, Simonoff E, Yuzda E, et al. Autism as a strongly genetic disorder: evidence from a British twin study. *Psychol Med*. 1995;25(1):63–77.
4. Hallmayer J, Cleveland S, Torres A, Phillips J, Cohen B, Torigoe T, et al. Genetic heritability and shared environmental factors among twin pairs with autism. *Arch Gen Psychiatry*. 2011;68(11):1095–102. doi:10.1001/archgenpsychiatry.2011.76.
5. Constantino JN, Zhang Y, Frazier T, Abbacchi AM, Law P. Sibling recurrence and the genetic epidemiology of autism. *Am J Psychiatry*. 2010;167(11):1349–56. doi:10.1176/appi.ajp.2010.09101470.
6. Ozonoff S, Young GS, Carter A, Messinger D, Yirmiya N, Zwaigenbaum L, et al. Recurrence risk for autism spectrum disorders: a baby siblings research consortium study. *Pediatrics*. 2011;128(3):e488–95. doi:10.1542/peds.2010-2825.
7. Iossifov I, Ronemus M, Levy D, Wang Z, Hakker I, Rosenbaum J, et al. De novo gene disruptions in children on the autistic spectrum. *Neuron*. 2012;74(2):285–99. doi:10.1016/j.neuron.2012.04.009.
8. Levy D, Ronemus M, Yamrom B, Lee YH, Leotta A, Kendall J, et al. Rare de novo and transmitted copy-number variation in autistic spectrum disorders. *Neuron*. 2011;70(5):886–97. doi:10.1016/j.neuron.2011.05.015.
9. Neale BM, Kou Y, Liu L, Ma'ayan A, Samocha KE, Sabo A, et al. Patterns and rates of exonic de novo mutations in autism spectrum disorders. *Nature*. 2012;485(7397):242–5. doi:10.1038/nature11011.
10. O'Roak BJ, Vives L, Girirajan S, Karakoc E, Krumm N, Coe BP, et al. Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations. *Nature*. 2012;485(7397):246–50. doi:10.1038/nature10989.
11. Pinto D, Pagnamenta AT, Klei L, Anney R, Merico D, Regan R, et al. Functional impact of global rare copy number variation in autism spectrum disorders. *Nature*. 2010;466(7304):368–72. doi:10.1038/nature09146.
12. Sanders SJ, Ercan-Sencicek AG, Hus V, Luo R, Murtha MT, Moreno-De-Luca D, et al. Multiple recurrent de novo CNVs, including duplications of the 7q11.23 Williams syndrome region, are strongly associated with autism. *Neuron*. 2011;70(5):863–85. doi:10.1016/j.neuron.2011.05.002.
13. Sanders SJ, Murtha MT, Gupta AR, Murdoch JD, Raubeson MJ, Willsey AJ, et al. De novo mutations revealed by whole-exome sequencing are strongly associated with autism. *Nature*. 2012;485(7397):237–41. doi:10.1038/nature10945.
14. Sebat J, Lakshmi B, Malhotra D, Troge J, Lese-Martin C, Walsh T, et al. Strong association of de novo copy number mutations with autism. *Science*. 2007;316(5823):445–9. doi:10.1126/science.1138659.
15. Gaugler T, Klei L, Sanders SJ, Bodea CA, Goldberg AP, Lee AB, et al. Most genetic risk for autism resides with common variation. *Nat genet*. 2014. doi:10.1038/ng.3039.
16. Stein JL, Parikshak NN, Geschwind DH. Rare inherited variation in autism: beginning to see the forest and a few trees. *Neuron*. 2013;77(2):209–11. doi:10.1016/j.neuron.2013.01.010.
17. Anney R, Klei L, Pinto D, Regan R, Conroy J, Magalhaes TR, et al. A genome-wide scan for common alleles affecting risk for autism. *Hum Mol Genet*. 2010;19(20):4072–82. doi:10.1093/hmg/ddq307.

18. Cantor RM, Kono N, Duvall JA, Alvarez-Retuerto A, Stone JL, Alarcon M, et al. Replication of autism linkage: fine-mapping peak at 17q21. *Am J Hum Genet*. 2005;76(6):1050–6. doi:10.1086/430278.
19. Liu J, Nyholt DR, Magnussen P, Parano E, Pavone P, Geschwind D, et al. A genomewide screen for autism susceptibility loci. *Am J Hum Genet*. 2001;69(2):327–40.
20. McCauley JL, Li C, Jiang L, Olson LM, Crockett G, Gainer K, et al. Genome-wide and ordered-subset linkage analyses provide support for autism loci on 17q and 19p with evidence of phenotypic and interlocus genetic correlates. *BMC Med Genet*. 2005;6:1. doi:10.1186/1471-2350-6-1.
21. Stone JL, Merriman B, Cantor RM, Yonan AL, Gilliam TC, Geschwind DH, et al. Evidence for sex-specific risk alleles in autism spectrum disorder. *Am J Hum Genet*. 2004;75(6):1117–23. doi:10.1086/426034.
22. Szatmari P, Paterson AD, Zwaigenbaum L, Roberts W, Brian J, Liu XQ, et al. Mapping autism risk loci using genetic linkage and chromosomal rearrangements. *Nat Genet*. 2007;39(3):319–28. doi:10.1038/ng1985.
23. Wang K, Zhang H, Ma D, Bucan M, Glessner JT, Abrahams BS, et al. Common genetic variants on 5p14.1 associate with autism spectrum disorders. *Nature*. 2009;459(7246):528–33. doi:10.1038/nature07999.
24. Weiss LA, Arking DE, Daly MJ, Chakravarti A. A genome-wide linkage and association scan reveals novel loci for autism. *Nature*. 2009;461(7265):802–8. doi:10.1038/nature08490.
25. Werling DM, Lowe JK, Luo R, Cantor RM, Geschwind DH. Replication of linkage at chromosome 20p13 and identification of suggestive sex-differential risk loci for autism spectrum disorder. *Molecular autism*. 2014;5(1):13. doi:10.1186/2040-2392-5-13.
26. Yonan AL, Alarcon M, Cheng R, Magnusson PK, Spence SJ, Palmer AA, et al. A genomewide screen of 345 families for autism-susceptibility loci. *Am J Hum Genet*. 2003;73(4):886–97.
27. Ronemus M, Iossifov I, Levy D, Wigler M. The role of de novo mutations in the genetics of autism spectrum disorders. *Nat Rev Genet*. 2014;15(2):133–41. doi:10.1038/nrg3585.
28. Zhao X, Leotta A, Kustanovich V, Lajonchere C, Geschwind DH, Law K, et al. A unified genetic theory for sporadic and inherited autism. *Proc Natl Acad Sci U S A*. 2007;104(31):12831–6. doi:10.1073/pnas.0705803104.
29. Reich R, Cloninger CR, Guze SB. The multifactorial model of disease transmission: I. Description of the model and its use in psychiatry. *Br J Psychiatry J Ment Sci*. 1975;127:1–10.
30. Tsai L, Stewart MA, August G. Implication of sex differences in the familial transmission of infantile autism. *J Autism Dev Disord*. 1981;11(2):165–73.
31. Werling DM, Geschwind DH. Sex differences in autism spectrum disorders. *Curr Opin Neurol*. 2013;26(2):146–53. doi:10.1097/WCO.0b013e32835ee548.
32. Carter CO, Evans KA. Inheritance of congenital pyloric stenosis. *J Med Genet*. 1969;6(3):233–54.
33. Robinson EB, Lichtenstein P, Anckarsater H, Hapke F, Ronald A. Examining and interpreting the female protective effect against autistic behavior. *Proc Natl Acad Sci U S A*. 2013;110(13):5258–62. doi:10.1073/pnas.1211070110.
34. Jacquemont S, Coe BP, Hersch M, Duyzend MH, Krumm N, Bergmann S, et al. A higher mutational burden in females supports a “female protective model” in neurodevelopmental disorders. *Am J Hum Genet*. 2014;94(3):415–25. doi:10.1016/j.ajhg.2014.02.001.
35. Lajonchere CM, Consortium A. Changing the landscape of autism research: the autism genetic resource exchange. *Neuron*. 2010;68(2):187–91. doi:10.1016/j.neuron.2010.10.009.
36. Geschwind DH, Sowiński J, Lord C, Iversen P, Shestack J, Jones P, et al. The autism genetic resource exchange: a resource for the study of autism and related neuropsychiatric conditions. *Am J Hum Genet*. 2001;69(2):463–6. doi:10.1086/321292.
37. Volkmar FR, Szatmari P, Sparrow SS. Sex differences in pervasive developmental disorders. *J Autism Dev Disord*. 1993;23(4):579–91.
38. Yeargin-Allsopp M, Rice C, Karapurkar T, Doernberg N, Boyle C, Murphy C. Prevalence of autism in a US metropolitan area. *JAMA*. 2003;289(1):49–55.
39. Sparrow SS, Balla DA, Cicchetti DV. Vineland adaptive behavior scales: interview edition, survey form manual. Circle Pines, MN: American Guidance Service; 1984.
40. Cheslack-Postava K, Liu K, Bearman PS. Closely spaced pregnancies are associated with increased odds of autism in California sibling births. *Pediatrics*. 2011;127(2):246–53. doi:10.1542/peds. 2010-2371.
41. Gunnes N, Suren P, Bresnahan M, Hornig M, Lie KK, Lipkin WI, et al. Interpregnancy interval and risk of autistic disorder. *Epidemiology*. 2013;24(6):906–12. doi:10.1097/01.ede.0000434435.52506.f5.
42. Risch N, Hoffmann TJ, Anderson M, Croen LA, Grether JK, Windham GC. Familial recurrence of autism spectrum disorder: evaluating genetic and environmental contributions. *Am J Psychiatry*. 2014. doi:10.1176/appi.ajp.2014.13101359.
43. Kong A, Frigge ML, Masson G, Besenbacher S, Sulem P, Magnusson G, et al. Rate of de novo mutations and the importance of father's age to disease risk. *Nature*. 2012;488(7412):471–5. doi:10.1038/nature11396.
44. Banach R, Thompson A, Szatmari P, Goldberg J, Tuff L, Zwaigenbaum L, et al. Brief report: relationship between non-verbal IQ and gender in autism. *J Autism Dev Disord*. 2009;39(1):188–93. doi:10.1007/s10803-008-0612-4.
45. Fombonne E. The epidemiology of autism: a review. *Psychol Med*. 1999;29(4):769–86.
46. Dunn LM, Bulheller S, Häcker H. Peabody picture vocabulary test. Circle Pines, MN: American Guidance Service; 1965.
47. Raven JC, De Lemos MM. Standard progressive matrices. London: Lewis; 1958.
48. Constantino JN. Social responsiveness scale manual. Western Psychological Services: Los Angeles; 2012.
49. Sato D, Lionel AC, Leblond CS, Prasad A, Pinto D, Walker S, et al. SHANK1 deletions in males with autism spectrum disorder. *Am J Hum Genet*. 2012;90(5):879–87. doi:10.1016/j.ajhg.2012.03.017.
50. Tropeano M, Ahn JW, Dobson RJ, Breen G, Rucker J, Dixit A, et al. Male-biased autosomal effect of 16p13.11 copy number variation in neurodevelopmental disorders. *PLoS One*. 2013;8(4):e61365. doi:10.1371/journal.pone.0061365.
51. Conde-Agudelo A, Rosas-Bermudez A, Castano F, Norton MH. Effects of birth spacing on maternal, perinatal, infant, and child health: a systematic review of causal mechanisms. *Stud Fam Plann*. 2012;43(2):93–114.
52. Gronborg TK, Schendel DE, Parner ET. Recurrence of autism spectrum disorders in full- and half-siblings and trends over time: a population-based cohort study. *JAMA pediatrics*. 2013;167(10):947–53. doi:10.1001/jamapediatrics.2013.2259.
53. Sandin S, Lichtenstein P, Kuja-Halkola R, Larsson H, Hultman CM, Reichenberg A. The familial risk of autism. *JAMA*. 2014;311(17):1770–7. doi:10.1001/jama.2014.4144.
54. Giarelli E, Wiggins LD, Rice CE, Levy SE, Kirby RS, Pinto-Martin J, et al. Sex differences in the evaluation and diagnosis of autism spectrum disorders among children. *Disabil Health J*. 2010;3(2):107–16. doi:10.1016/j.dhjo.2009.07.001.
55. Dworzynski K, Ronald A, Bolton P, Hapke F. How different are girls and boys above and below the diagnostic threshold for autism spectrum disorders? *J Am Acad Child Adolesc Psychiatry*. 2012;51(8):788–97. doi:10.1016/j.jaac.2012.05.018.
56. Kim YS, Leventhal BL, Koh YJ, Fombonne E, Laska E, Lim EC, et al. Prevalence of autism spectrum disorders in a total population sample. *Am J Psychiatry*. 2011;168(9):904–12. doi:10.1176/appi.ajp.2011.10101532.
57. Mattila ML, Kielen M, Linna SL, Jussila K, Ebeling H, Bloigu R, et al. Autism spectrum disorders according to DSM-IV-TR and comparison with DSM-5 draft criteria: an epidemiological study. *J Am Acad Child Adolesc Psychiatry*. 2011;50(6):583–92. doi:10.1016/j.jaac.2011.04.001. e511.
58. Zwaigenbaum L, Bryson SE, Szatmari P, Brian J, Smith IM, Roberts W, et al. Sex differences in children with autism spectrum disorder identified within a high-risk infant cohort. *J Autism Dev Disord*. 2012;42(12):2585–96. doi:10.1007/s10803-012-1515-y.